Fig. 1

(Prior Art)

24 — dma
12 — switch fabric
14 — cpu
22 — cpu
16 — ram
20 — i/o
18 — ethernet
chip SoC 10

Fig. 2

48 — ext i/f
32
34
38 — cpu
bridge
dma
46
timer
44
36
peripheral
42
i/o — 40
30

How Rings assign address space

Step1: align incoming address to self (to some power of 2)

Step2: assign the result to self address

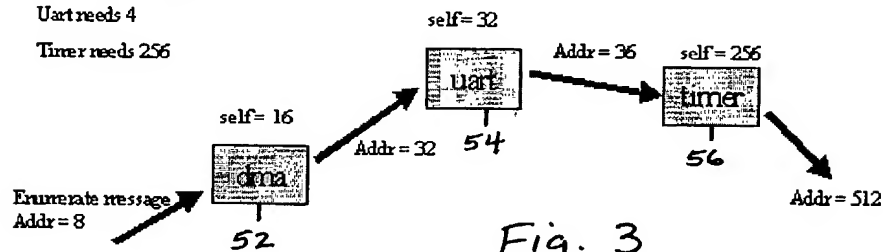Step3: next_addr = self_addr + self_addr_space; // number of register used locally

Step4: send down next_addr

Example:

Dma needs 16 addrs

Uart needs 4

Timer needs 256

self = 32

Addr = 36   self = 256

self = 16

Addr = 32   54

Enumerate message
Addr = 8

52

56

Addr = 512

Fig. 3

Fig. 4

60   62

If the delay between "clk1" and "clk2" greater then the delay from Q to d of second flipflop, we have a race on our meaning right hand flipflop will sample the data of Q a whole clock period early.

clk1   clk2

Fig. 5

72

74   76

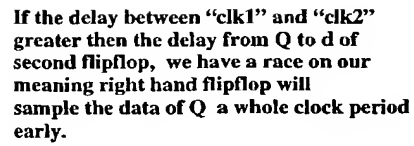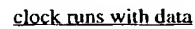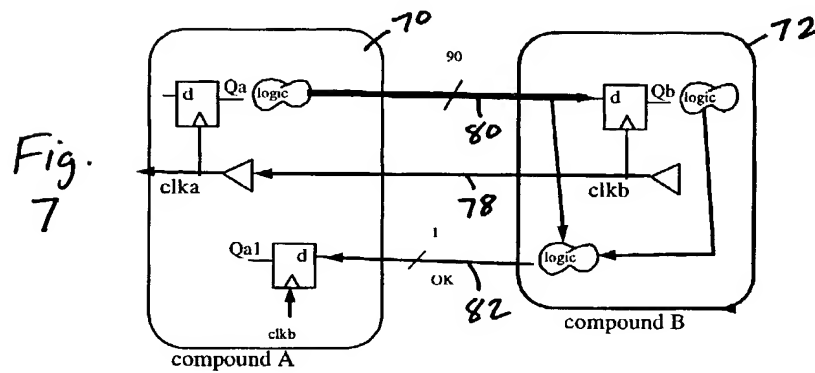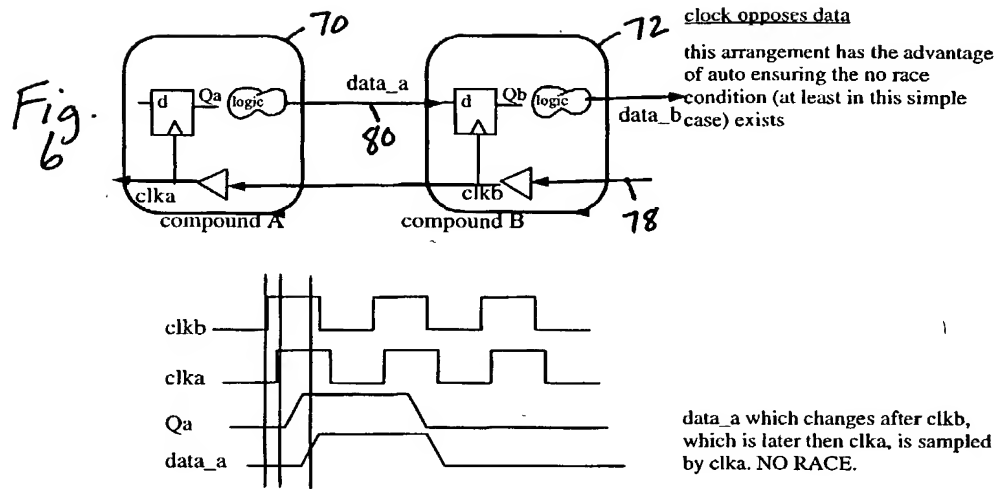clk   66

compound A   70   compound B

64

clock runs with data

the problem is possible race. However, we control the logic on each flipflop leaving the compound, because it is always the same standart ring-interface module. we can ensure, that the delay will be at least enough. And more importantly easily checked after layout.

Fig. 6

Fig. 6 diagram:

70 compound A — d, Qa, logic — data_a 80 — 72 compound B — d, Qb, logic — data_b

clka — clkb — 78

clock opposes data

this arrangement has the advantage of auto ensuring the no race condition (at least in this simple case) exists

Timing diagram:
clkb
clka
Qa
data_a

data_a which changes after clkb, which is later then clka, is sampled by clka. NO RACE.

Fig. 7

Fig. 7 diagram:

70 compound A — d, Qa, logic — 90 — 80 — 72 compound B — d, Qb, logic

clka — 78 — clkb

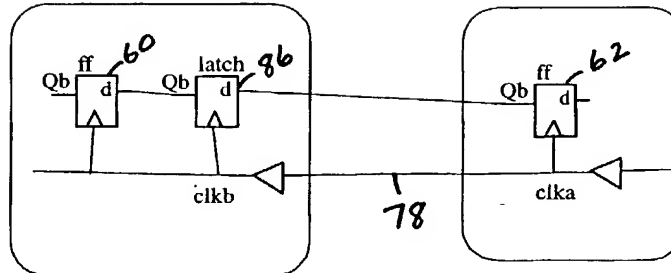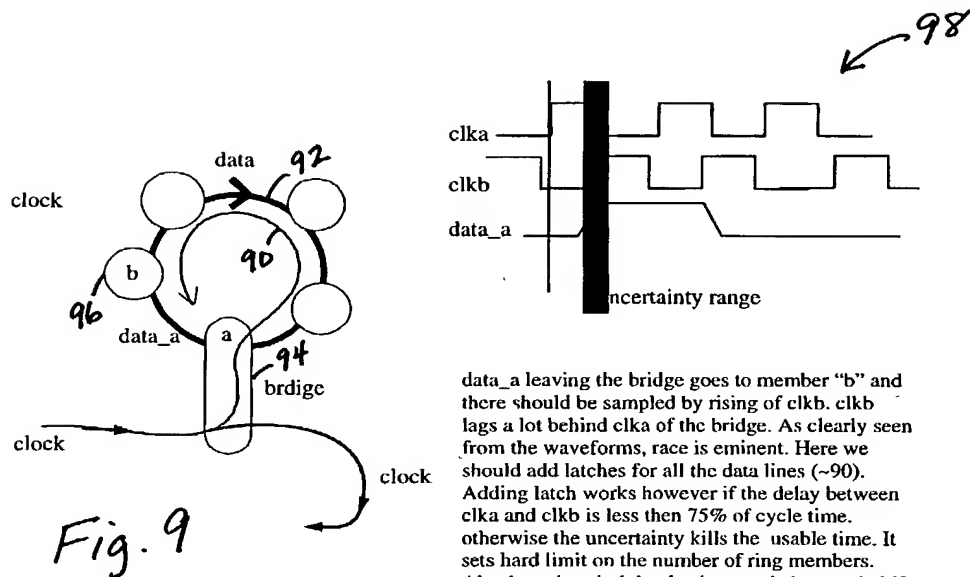Qa1 — d — 1 OK — logic — 82

clkb

compound A

compound B

Fig. 8

Fig. 9

data_a leaving the bridge goes to member "b" and there should be sampled by rising of clkb. clkb lags a lot behind clka of the bridge. As clearly seen from the waveforms, race is eminent. Here we should add latches for all the data lines (~90). Adding latch works however if the delay between clka and clkb is less then 75% of cycle time. otherwise the uncertainty kills the usable time. It sets hard limit on the number of ring members. Also keep in mind that latches needed on each OK signal between members of the ring
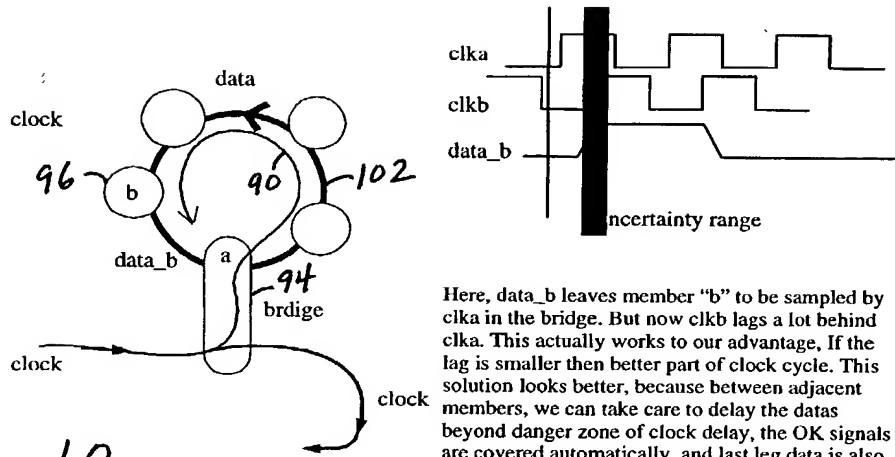
clock

data

96 ~

b

90

~102

data_b

a

~94

brdige

clock

clock

clock

Fig. 10

clka

clkb

data_b

Uncertainty range

Here, data_b leaves member "b" to be sampled by clka in the bridge. But now clkb lags a lot behind clka. This actually works to our advantage. If the lag is smaller then better part of clock cycle. This solution looks better, because between adjacent members, we can take care to delay the datas beyond danger zone of clock delay, the OK signals are covered automatically, and last leg data is also covered. The only signal not safe is the OK from bridge to "b" member. It will need a latch in "b".
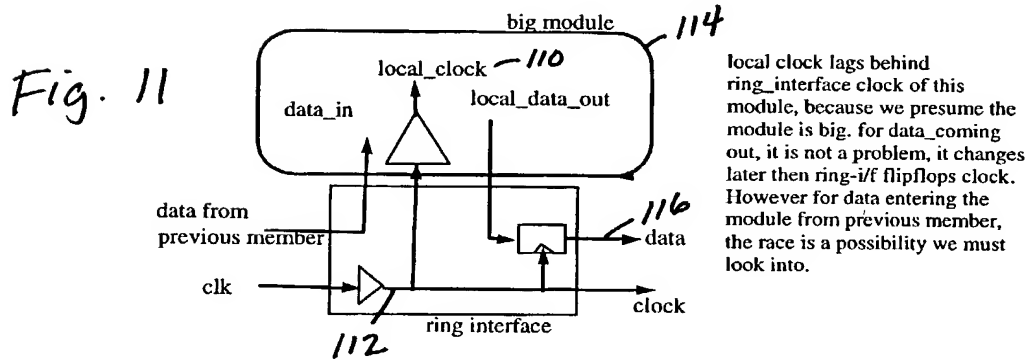
Fig. 11

big module ~114

local_clock ~110

data_in

local_data_out

data from previous member

~116

data

clk

clock

112

ring interface

local clock lags behind ring_interface clock of this module, because we presume the module is big. for data_coming out, it is not a problem, it changes later then ring-i/f flipflops clock. However for data entering the module from previous member, the race is a possibility we must look into.
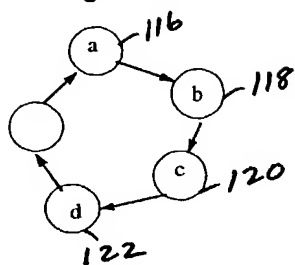
Fig. 12

a — 116
b — 118
c — 120
d
122

if module "a" sends a message to module "b", ring works fine. However if most of the traffic is from "c" to "b", this is more expensive in terms of latency.

Another problem is "peak latency". Suppose that , "a" transmits mostly to "d" and "b" mostly to "c" In this case communication between "b" and "c" suffers degradation in case that peak traffic coincide.
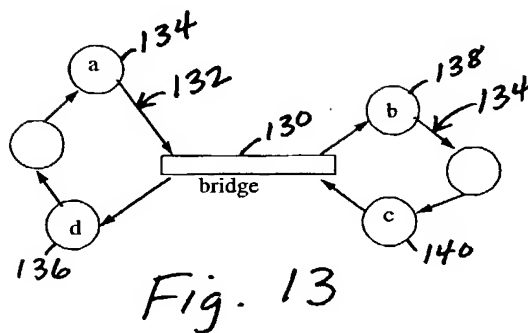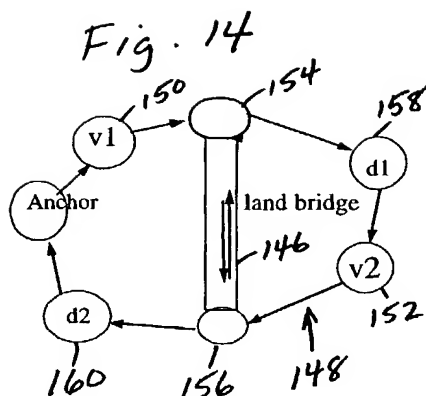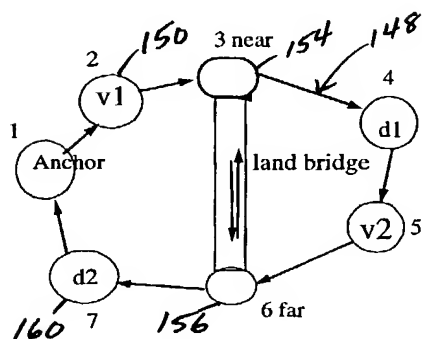

a — 134
— 132
b — 138
— 134
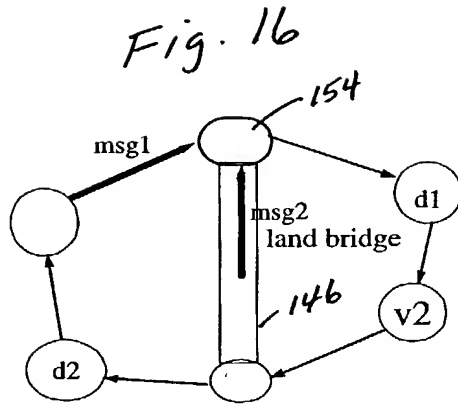— 130
bridge
d
136
c
140

Fig. 13

Fig. 14

Land bridge gets its name from the fact that it is a luxury. It spans across connected modules. The idea is simple. When V2 sends message to D1 it gets to one side of the bridge. This side analyzes the destination address and by some magic (explained later) decides to short-cut the path. The message re-appears at the other end of the bridge and gets fast to D1. By same magic, message from V1 to D2 get bypassed also. message from V1 to D1 is treated directly.

Enumeration is started by "Anchor" which assigns address=1 to itself. results of enumeration are labels 1 to 7. land bridge gets two addresses , as if it were not one module. there is "near" end, that got enumeration label "3", and the "far" end marked 6.

Fig. 15

Fig. 16

msg1

_154_
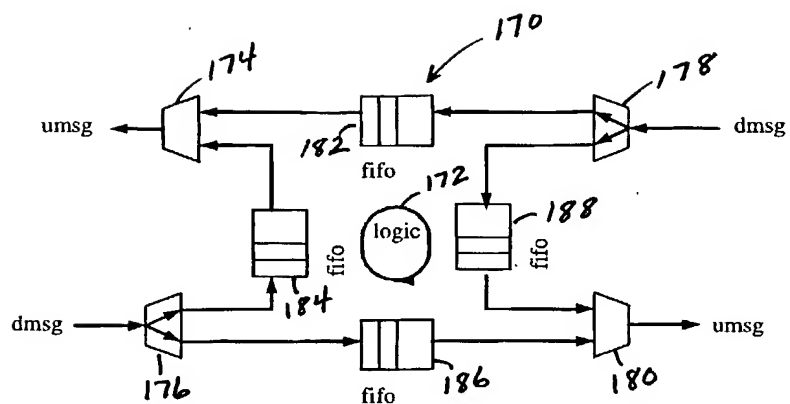
d1

msg2
land bridge

_146_

v2

d2

msg1 and msg2 arrive at the same time. the bridge end must make a decision which message to forward first.

It can be shown that unwise decision can lead to freezout, deadlock and option price dropping to 5$.

Therefore MSG2 gets the priority.

Fig. 17

Fig. 18

a

anchor

192

d

194

near | far
bridge

b

c

Bridge takes responcibility for strays, but only at the "far" end. During enumeration, bridge is "polarized" to have near and far end. Near is the end first struck by enumeration message.

So we have exactly one enforcer for each ring.

3 near

2

v1

1

Anchor

202

198

land bridge

4

d1

v2 10

196

d2

12

200

11 far

Fig. 19

In land bridge ring, the situation is trickier. If V2 send message to address==5. The land bridge divert at 11/far end. it will re-appear at 3 and start cycling forever.

We have to define an algorithm that will take care of all cases.

Luckily there is a way.

Land Bridge deals only with messages arriving at the far end and being diverted. It marks and monitors only those. Messages arriving at near end, keep their markings. Messages at fdar end going through, are left alone.

Fig. 20

2
v1

3 near

4
d1

1
Anch

210

land bridge

land bridge

208

land bridge

d2

12

11 far

v2 10

206

Fig. 21

2
v1

3 near

4
d1

1
Anchor

214

land bridge

216

d2

12

11 far

v2 10

212

218

220

a

land bridge

224

b

bridge

d

c

Fig. 22

Fig. 23

ring chip
no scan

ring chip
scan added

scan
insert

230

Fig. 24

scan insert module

ring out   ring in

no
mux   232
yes

230   scan mode

tap the results

out pads

236

insert scan

in pads   234

Fig. 25

rings
bus
crossection

240   8   msg_type

242   20   msg_addr

246   64   msg_data(63:8) not used for scan

8   msg_data(7:0) , used for scan chains

248   clock

scan_mux

250

during the first clock, OK remains active, when type is of msgA. It means that on the next clock, memberA may send new message. memberA uses this ok to send msgB on the next clock. msgB gets stuck for a clock because OK goes inactive. It goes inactive because the fifo in memberB is full. One clock later, the fifo has a free entry, so OK returns to 1 and type returns to idle next clock. return to idle could also be change to next message, if there was one.
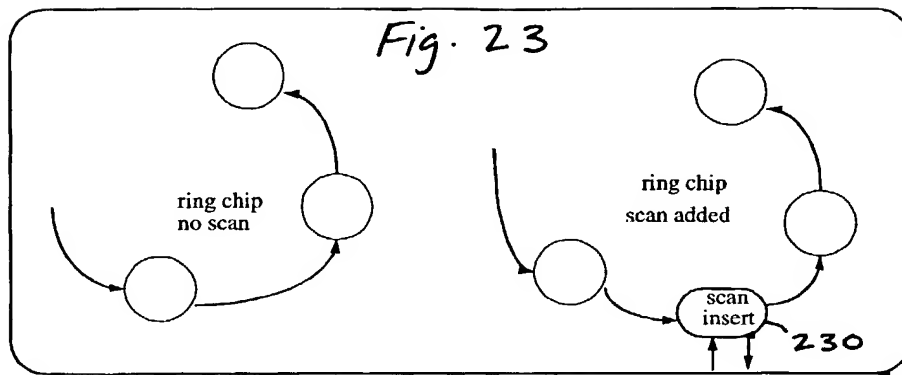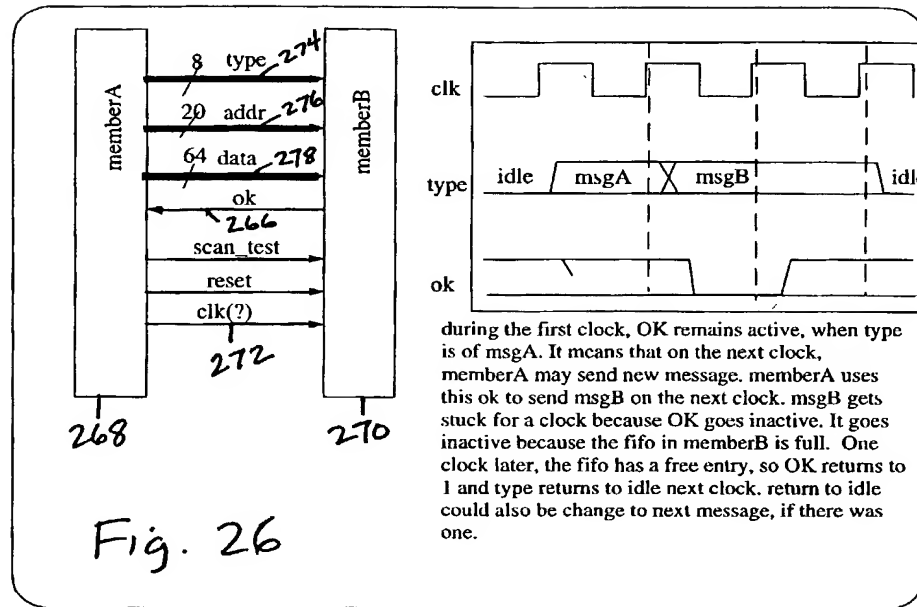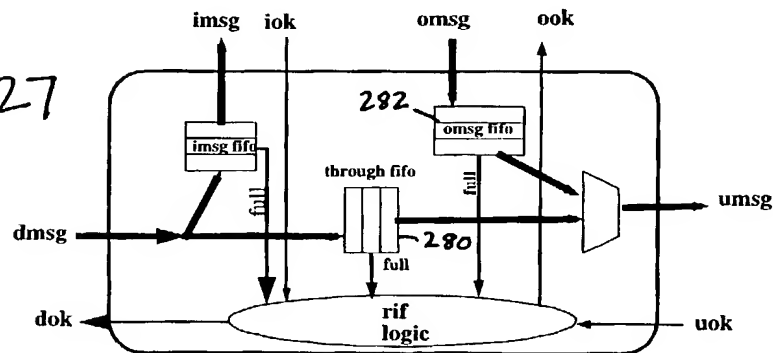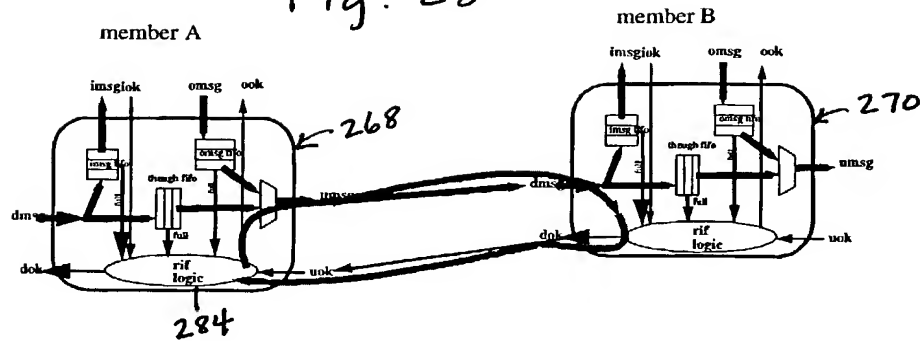
Fig. 26

Fig. 27

Fig. 28

member A
member B

268

270

284

The incoming messages are examined first to see if it is supervisor or work/program. Work/program messages have address field. We check if it is our address. Since we know that our address is aligned to our power of 2, The address mask (named split mask) causes only certain number if upper bits to be compared. The lower part of the address is passed inside as internal address. The upper bits are compared against self-address register. This register gets its value during enumeration protocol. The lower part of this register is always masked,. Hopefully synthesis will delete the unused bits implementation.

19   290   0

incoming
address

294

292
address
split mask

self address
register

comparator

dont care part
of self-address

296

ours/through

part of the address
that enters the member

Fig. 29

Fig. 30

Member

Rif_I_options[5:0]
Rif_I_addr[*:0]
Rif_I_datal/h[31:0]

Rif_I_write
Rif_I_read

Rif_I_ok

Rif_I_clock

RIF   * = address_space

Activation register

300

Fig. 31

Fig. 32

300

Member

Rif_I_*   **Rif_***   **Rif_o_***   **module_id**

RIF   Address Space = 7

| Activation register |

Fig. 33   300

Fig. 34

the second land bridge solves most traffic problems, but adds 4 clocks in the overall ring length. This is not a big problem because no message should travel the whole perimiter.

Fig. 35

The utopia interface is forced into mode that communicates in messages, not cells. We using the I/O and maybe some of the logic.

Fig. 36

Application Specific Accelerators	Internal Memory 352 Fast, Unified, Multi-port	System Expansion Area
CRC Encryption Table Lookup Hashing 358	Doorbell Vobla 354 Network Processor	CPU (PP) DMA Smart FIFO Ext. mem I/F
	Peripheral Expansion Area 356 Enet, ATM, Uart, USB, Serials	360

350

Fig. 37

Fig. 38

Current Task `Task_X`
(CTID)

To Memory

Next Task `Task_Y`
(NTID)

Logic

Logic

Write Queue

Memory

Preload regs of Task_Y

Destination

Tag

Active
(Active Reg. File)

Shadow1
(Active Shadow)

Shadow2
(Preload Shadow)

← 390

hit/miss
Source

H   M
Mux

Source Operand
of Task_X

– – – – – – – – – After a task switch – – – – – – – – – –

Current Task `Task_Y`
(CTID)

To Memory

Next Task `Task_Z`
(NTID)

Logic

Logic

Write Queue

Memory

Preload regs of Task_Z

Destination

Tag

Active
(Active Reg. File)

Shadow1
(Active Shadow)

Shadow2
(Preload Shadow)

← 392

hit/miss
of Source

H   M
Mux

Source Operand
of Task_Y

Fig. 39

PP ◄──────── 408 ──────► External memory

Other data ◄──────────► NP task chain

406

Internal memory

Internal memory

Peripheral Fifo

"The system"

Peripheral Fifo

402

"External world"

stream data in

stream data out

Fig. 40

420

32 registers of 32 bits per task ◄── General purpose registers

Special purpose registers ──► per task and global registers

A set of indications per task, which control task execution scheduling ◄── Doorbells

NP configuration registers ──► Initialized by the PP

An interface to adjacent resources ◄── Agent interface

Fast memory accessed by load/store instructions ◄── Internal memory

External memory ──► Big area accessed via a DMA interface

Fig. 41

R1 register:

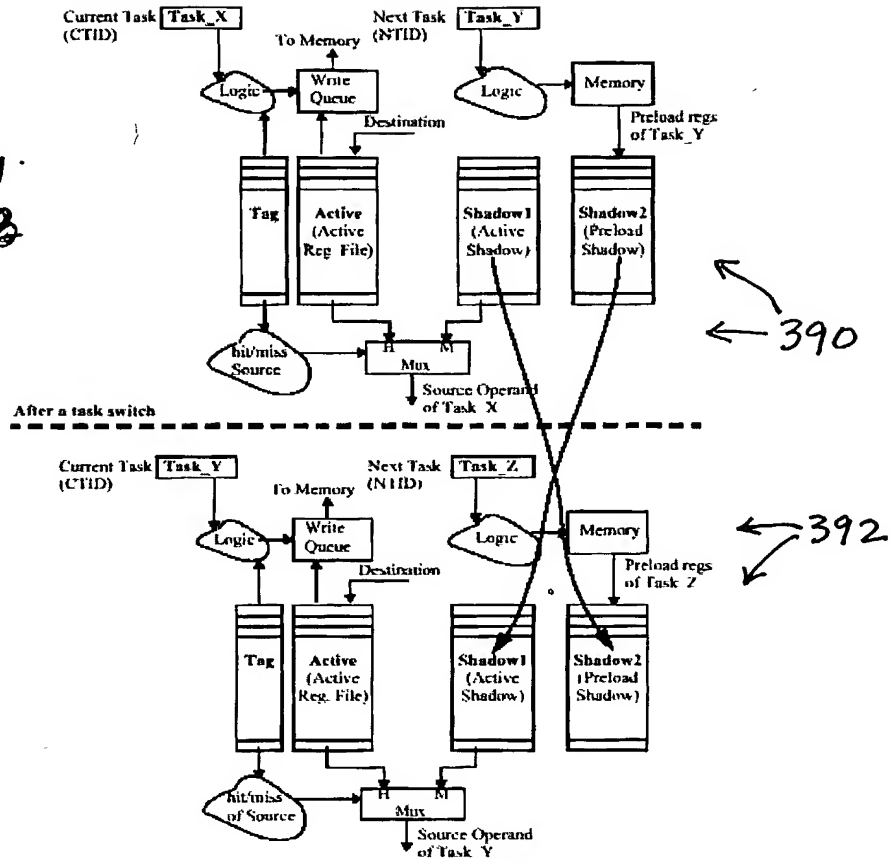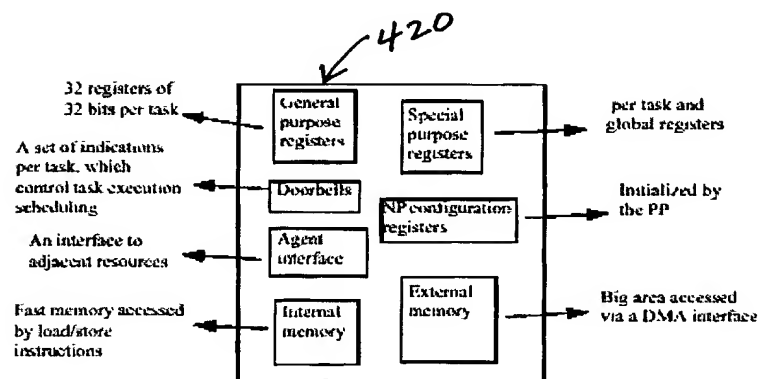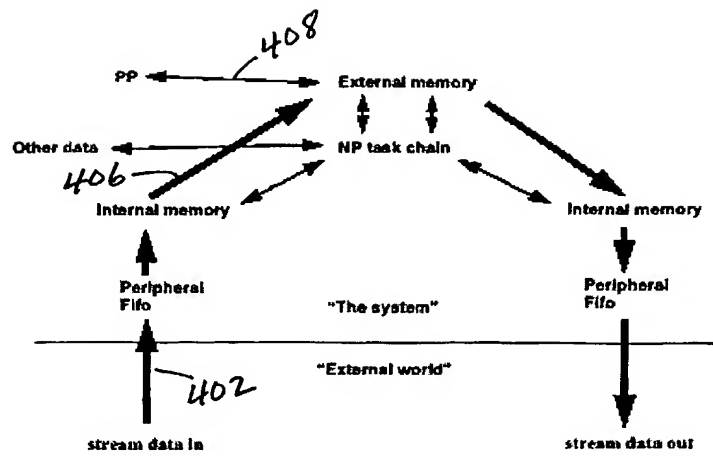| 1 1 2 3 2 2 2 2 2 2 1 1 1 1 1 3 1 1 1 9 8 ? 6 5 4 3 2 1 0 |
| L b 9 K ? 6 ? 4 3 2 1 0 9 8 ? ? 5 4 3 2 1 0 |

430

s - sticky bit
eq – equal/zero
lt - less then/negative
gt - greater then/positive
c - carry
mb - reflection of the RAM multi-reader busy indication.

Fig. 42

REFETCH SPR
(spr index = 0)

| 3 3 2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1 9 8 7 6 5 4 3 2 1 0 |
| 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 |

| NEXT REFETCH | REFETCH |

440

TASK SPR
(spr index = 1)

| 3 3 2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1 9 8 7 6 5 4 3 2 1 0 |
| 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 |

| DOOR BELL REQ | | | COUN T | | | 1 R | MASK | | N T V | NHD | | CHD |

442

TRAP SPR
(spr index = 2)

| 3 3 2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1 9 8 7 6 5 4 3 2 1 0 |
| 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 |

| | | | | | | | | | | | Y B | I B | D A B | P A B | T R A P | E I |

444

MINDEX SPR
(spr index = 3)

| 3 3 2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1 9 8 7 6 5 4 3 2 1 0 |
| 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 |

| | INDEX2 | | INDEX1 |

446

Frame structure
of an example
task type

450

common task data	
task fragment 1 data	task fragment 2 data	task fragment 3 data
data of all level2 functions	
level1 f1 data	level1 f2 data	level1 f3 data

r27

Fig. 43

- size of level0 frame part is different for each task type
- size of level2 frame part is constant
- size of level1 frame part is different per each task type

460   *Fig. 44*

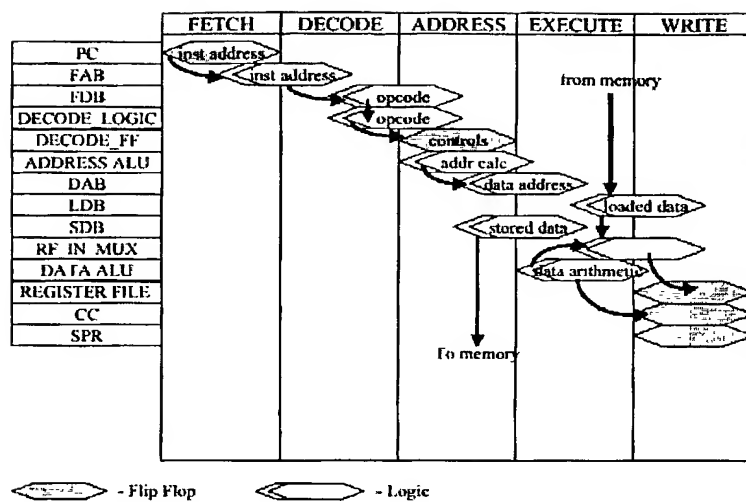|	FETCH	DECODE	ADDRESS	EXECUTE	WRITE
PC					
FAB					
FDB					
DECODE_LOGIC					
DECODE_FF					
ADDRESS ALU					
DAB					
LDB					
SDB					
RF_IN_MUX					
DATA ALU					
REGISTER FILE					
CC					
SPR					

480

Fig. 45

⬡ - Flip Flop   ⬡ - Logic

Fig. 46

500

Fig. 47

Fig. 48

Fig. 49

Fig. 50

Fig. 51

528


```
Vobla  agent interface   input_sel → main entry → output_sel → output message encoder   → type_out[7:0]
        stall_vobla                    shadow entry                                      → uddress_out[23:0]
                                                                                         → data_out[63:0]
                                                                                         ← reset
                                                                                         ← clk
                                                                                         ← ok2drive
```

560

562

Fig. 52

```
agent command          | opcode | options[9:0] | RA | AID[4:0] | RB/imm8 |
(AID=message_sender)
(option[6] =0)                                         RA+1
message_sender32 {0,options[5:0]| raw_data[31:0] | raw_address[23:0] | type[7:0] |
                        message data            address_of_destenation message type
```

```
agent command          | opcode | options[9:0] | RA | AID[4:0] | RB |
(AID=message_sender)
(option[6] =1)                                    RA+1
message_sender84 {1,options[5:0]| raw_data[63:0] | raw_address[23:0] | 10000000 |
                        message data            address_of_destenation message type
```

530

Fig. 53


```
doorbell   dm_set_mask 0
           dm_set_mask 1
           dm_set_mask 2
           dm_set_mask 3

           token control ↔ request entries X2 ← input message decoder   ← type_in[7:0]
                                                                          ← write_interface_address[23:0]
                                                                          ← write_interface_data[31:0]
                                                                          ← rit_base_address[7:0]

Vobla   agent interface
        stall_vobla
        current_task_id[5:0]       DMA context table → output message encoder → type_out[7:0]
                                                                                → address_out[23:0]
                                                                                → data_out[63:0]
                                                                                ← clk
                                                                                ← ok2drive
                                                                                ← reset
```

Fig. 54

agent command
(AID=dma agent) | opcode | options[10:0] | RA | AID[4:0] | RB/imm8 | ~576

RA+1

dma request entry | M | U | D | A | NA | L

modify address (OP9) | urgent (OP2) | dir (OP1) | autosend set (OP0) | long ack (OP3) | address (OP10)

dram address[31:0] | sram address[23:0] | {count[7:0]} | ~578

dram address | sram address | number of bytes to transfer or address modifyer

Fig. 55

last_in_transfer
last_in_frame
snoop
calculate_crc

multireader

multireaddata[31:0]
data size[2:0]
stall_multireader

register file | input buffer

tx_data_mux ← on_dmand

CRC data

Vobla

agent interface
stall_vobla

random number generator (TBD) | CRC 5,10,32 machines | checksum machine | bip16 machine

~520

clk
reset

Fig. 56

agent command
(AID=CRC agent) | opcode | options[9:0] | RA | AID[4:0] | RB | ~590

RA+1

type[2:0] | size[2:0] | G | S | O | DATA[63:32] | DATA[31:0] | RESIDUE

CRC type | data size | generate check | operation mode | overwrite residue

592

Fig. 57

526

Fig. 58

602

604

516

Fig. 59

Agent interface block (Vobla) with signals:
- agent interface
- v_task_switch
- v_set_doorbell_mask
- v_next_task_valid
- v_current_task_id[5:0]
- v_current_task_id valid
- v_next_task_id[5:0]
- v_doorbell_req[2:0]
- v_doorbell_mask[2:0]
- v_doorbell_count[1:0]
- v_urgent
- v_aset

Internal blocks: input message decoder, task mask registerfile, task requests register file and counters, TGMR, mask control logic, priority logic

Outputs:
- rif_i_write
- rif_i_doorbell_cs
- rif_i_addr[5:0]
- rif_i_data[4:0]
- dm_set_mask0
- dm_set_mask1
- dm_set_mask2
- dm_set_mask3
- rif_i_reset
- rif_i_clock

dma agent

612

Fig. 60

agent command (AID=doorbell agent)

| opcode | options[9:0] | RA | AID[4:0] | RB/imm8 |

RA+

SG — set/clear global (OP2)
CR — clear request (OP1)
CM — clear mask (OP0)
global task register mask

614

{0,0,0,0,0,mask_bit_index[2:0]} write mask
or
{0,0,0,0,1,req_bit_index[2:0]} write request
or
{1,0,0,0,0,count_value[2:0]} write counter
or
{0,1,0,0,0,0,0,0} write TGMR
or
{1,1,0,0,0,0,0,0,urgent_value} write urgent

Fig. 61

Fig. 62

Fig. 63

Trajan 1

Trajan 2

722

FPGA

Encryption — 734

DSP — 732

PCI
*(good also to
debug and insert
ring data)*

Fig. 64

Processor

Functional unit

External ring interface

Fig. 65

Memory

742

744

Encryption

748

Trajan chip | Memory port

CS

CS

Data

Addr

WR FIFO

DSP

746

RD FIFO

750

FIFO_RD

FIFO_WR

FPGA

PCI

740

740

760

Host #1 (PP)

Host #2 (DMA)

Memory controller

CS

ADDR

DATA

Ring interface

DATA

OK

Message sender

DATA

OK

Write request generator

Read request generator

WR FIFO used words counter

RD FIFO used words counter

FIFO_RD

FIFO_WR

Host #3 (Ring extender)

Trajan

Fig. 66

Fig. 67

mii

enet rx mac

setup registers:

doorbell, taskid and viscode
urgent viscode
header len and address
threshold to urgent

rx manager — 802

doorbells

fifo
ram

reader

data
header

setup

ring in

ring control

ring c

free		crc	ovrn	err	last	size	

rx status word

Fig. 68

mii

enet tx mac

setup registers:

doorbell, taskid and viscode
urgent viscode
send ahead address
threshold to transmit
threshold to urgent

fifo
ram

812

txmanager

setup

doorbell
free entry count
finished frames cou

ring in

ring control

ring c

Fig. 69

Control Plane

Signaling Protocols
Protocol Management
Exception Handling
System Control &
Configuration

Data Plane

Per/packet handling
Forwarding Decision
Classification
QoS Handling
Queuing
Scheduling
Formatting

830

- Protocol Processor

Memory interfaces

peripherals

Network Processor

Network Processor

834   836

Fig. 70

840

MGCP (SM)	MEGACO (SM)		OSPF (SM)	BGP-4 (SM)	RIP (SM)		MPLS (SM)	LDP CR-LDP RSVP-TE (SM)		UNI Signaling (SM)

UDP/TCP (SM)   IP (SM)   IP/ATM (SM)   ATM (SM)

MPLS (2.5*)

IP FWD (3*)

FRF.9 (2*)

FRF.5 (2*)

AAL-2 L366.1 (2*) AAL-2 L366.2 (3*) AAL-5 Switching (2*)

IPoA & PPPoA (2*)

AAL-2 L363.2 (2)

RFC 1490 (2*)

Circuit Emulation Services (CES) (2*)

RFC 1483 (2*)

802.1D Bridge (2*)

Frame Relay (2)

Cell Switch (2*)

MPLS (2.5*)

ML-PPP

AAL-0 (2) AAL-1 (2) AAL-3 (2) AAL-5 (2)

PPPoE (2*)

HDLC (2) Transparent (2)

ATM Low Layer Module (2)

Serial (1) PCM (1)

Utopia Interface (1)

Ethernet (1)

Fig. 71

Fig. 72

Fetch	Decode	Address	Execute	Write
Instruction fetch request	Instruction decode	Address calculation. Data access req.	Read source Registers Data execution	Write result into destination register

Fig. 73

◇ - Flip Flop   ◇ - Logic

Fig. 74

Calendar Wheel

Utopia

Fig. 75

- CBR – TDM Traffic AAL1
- VBR – Voice AAL2, RTP (SP), Streaming-Data
- UBR – Data oriented application
  - WEB browsing
  - FTP
  - Telnet
  - SMTP

Fig. 76

Fig. 77

Fig. 78

Fig. 79

Fig. 80

1000

	TDM Bus	ATM Bus		
SHDSL Line Card TDM <-> ATM				
T3/E3 Line Card ATM <-> ATM			MAN card ATM <-> ATM	ATM MAN
T1/E1 Line Card ATM/FR <-> ATM			MAN card ATM <-> IP	
Optical Interface ATM <-> ATM				
Ethernet Line Card TDM <-> ATM			Voice Gateway GR-303, V5.2 IP/ATM<->TDM	CLASS 5 Switch
xDSL Line Card ATM <-> ATM				

Fig. 81

1020

- CPP OAM
- CPP IP Termination
- IP over RFC 1483
- Eth over RFC 1483
- Eth over RFC 1483
- 802.1D Bridge
- CPP Spanning tree
- AAL-5
- AAL-5
- ATM Device driver / Utopia Interface
- ATM Device driver / Utopia Interface
- Ethernet

1030

Fig. 82

1040

Fig. 83

Fig. 84

1050

Fig. 85

Fig. 86

compare1 ~ 1402

branch to
label A if ~ 1404
cond. sat.

1400

compare2 ~ 1406

branch to
label B if ~ 1408
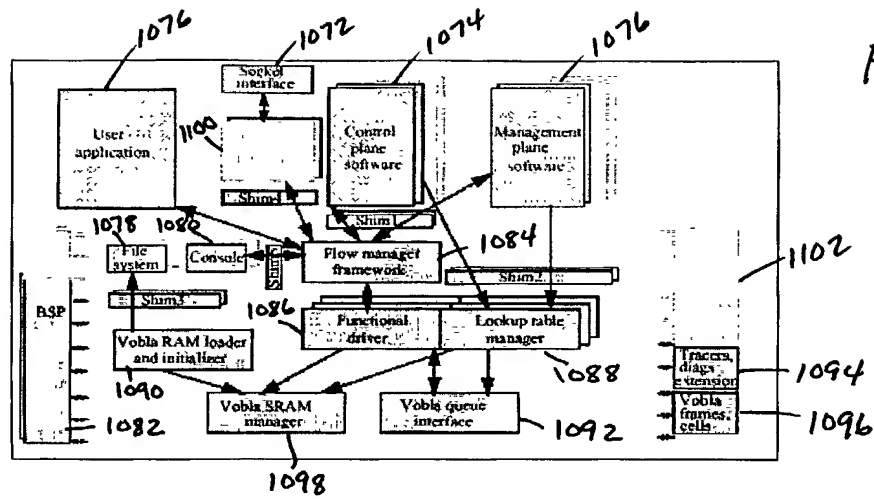cond. sat.

Compare 3 ~ 1410

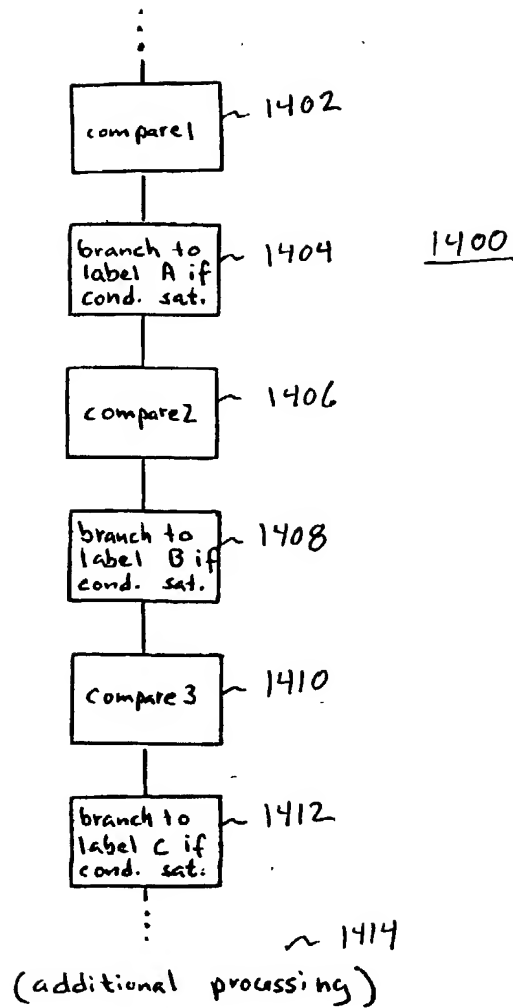branch to
label C if ~ 1412
cond. sat.

~ 1414
(additional processing)

FIG. 87
(PRIOR ART)

FIG. 88

FIG. 89

FIG. 90